

Strategic Evolution of ESE Data Systems

Lifecycle Study Breakout Report

Matt Schwaller
NASA/GSFC
schwaller@gsfc.nasa.gov

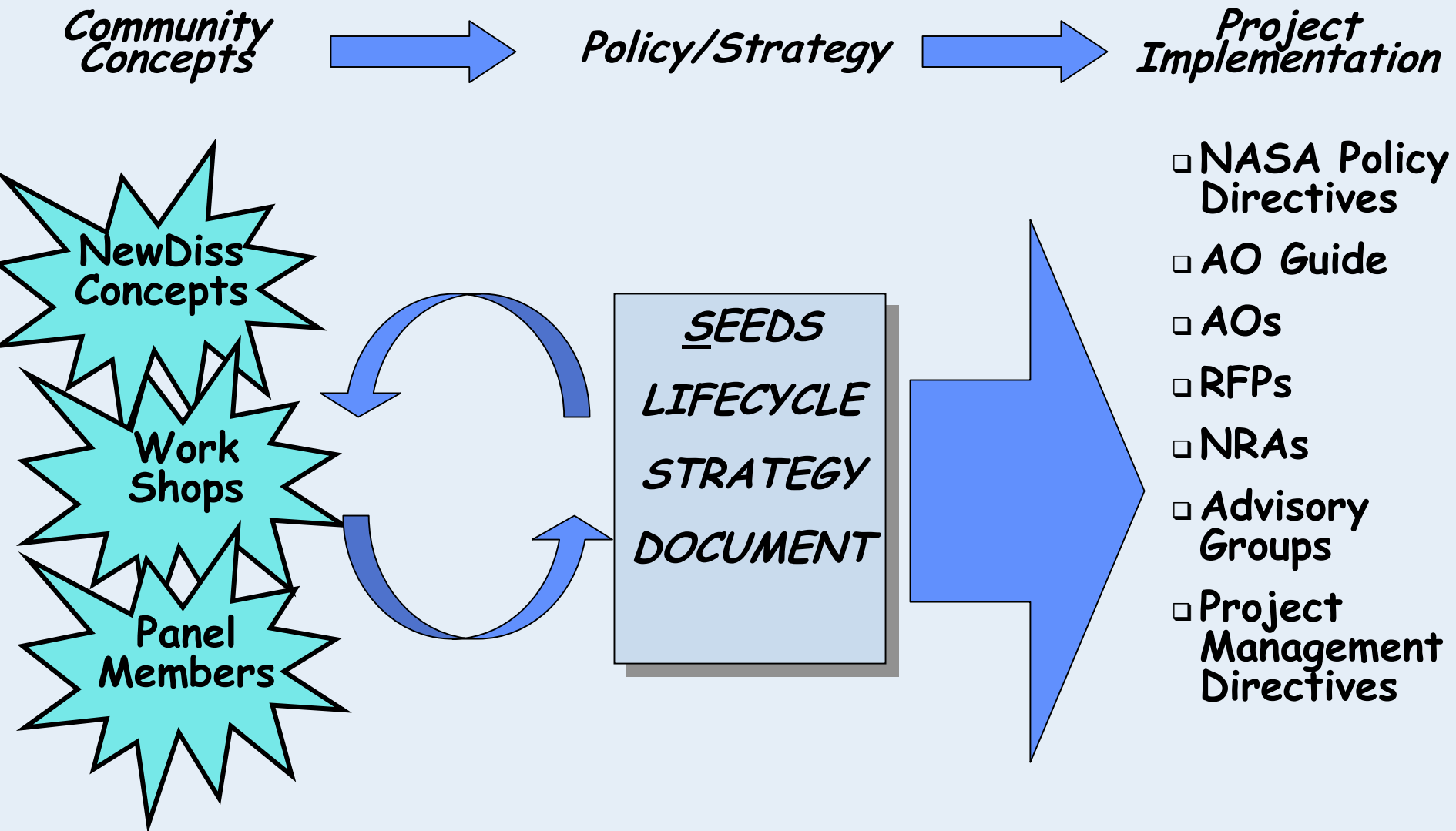
SEEDS Second Public Workshop
June 17-19, 2002

Yesterday's Agenda



- 1:00 (30 minutes) Introductions, study team results (Schwaller)
...and review the agenda proposed here...
- 1:30 (90 minutes) Panel Comments + Discussion
 - Jon Christopherson -- Spacecraft/Pre-launch
 - Robert Wolfe -- Product Generation
 - Steve Kempler -- Active Archive
 - Ben Watkins -- LTA
- 3:00 (60 minutes) Break into 4 sub-groups to review & comment on the report
- 4:00 (60 minutes) Report comments/findings group by group
- 5:00 Final words & adjourn

Objectives Restated



Results-Mission & Product Generation



Issues/Recommendations:

- ❑ Review the CCSDS reference model for **Open Archival Information System (OASIS)** and at least sync up with the terminology
- ❑ **"Data-buys" or any proprietary mission** needs a time-period to be defined after which the data becomes the property of NASA
- ❑ **Intellectual property rights** issues need to be addressed
 - PI institution needs to commit/agree to make all archive data products available (where possible) to the public (or a sunset clause needs to be defined)
 - Or, at least identify which tools and products were used, and identify how the data may be accessed
- ❑ Relevance of **data quality act?** (Watermark, provenance issues, reproducibility of data, peer review, integrity of data and supporting information)
- ❑ Issue of **keeping, discarding data**: CERES doesn't throw anything away ... MODIS does discard old versions
- ❑ Preservation of "community knowledge": need early involvement of mission and product generation communities in defining the content and LOS of the archive
 - Involvement of these people in advisory groups
- ❑ Need to **record production flows**, production histories, PGE-versions, etc.
- ❑ **Cost model**: need to parameterize LOS, levels of documentation -- what is the level of documentation necessary?
 - **Need to bring the lifecycle and cost teams together**

Results-Active & Permanent Archives



Issues

- Defined relationships in **data models** from various archives
- Acquisition of **documentation**
- What data is **serviced** by permanent archive (PA)
- What is the highest level **documentation** needed to go with data products to PA
- **Resources** for maintaining all data that is desired to keep
- No consistent policy/implementation for archiving mission/instrument **documentation**
- **Transfer electronically** only

Recommendations

- **Develop specific ICDs** described in detail information about the data (formats, volumes) between data producers and active archive; active archive and permanent archive
- Develop a **prioritization process** for selecting data/selecting **levels of service (LOS)**, involving user groups
- Collect for archival mission/instrument **documentation**
- Digitized **documentation** only to PA
- Development of **Project Data Management Plan (PDMP)** for each mission to include end to end plans for data archive
- Pre-mission review of **documentation** and by milestone event
- Provide language to AO's, etc. that requires appropriate level of **documentation**
- Study distributed permanent archives. Including
 - A cost/time table for achieving this
 - LOS at each site (per product)
 - Archive in place

Recommended Next Step



□ Immediate: Revise the guidelines document with ...

- Inputs from the 4 groups
- Revise scope, contents and possibilities in each area
- Incorporate today's issues/recommendations Near term: integrate study results with other study results
- Especially, parameterization of the SEEDS cost study